

2570  
09/6

# ENTERED



OIKE

## RAW SEQUENCE LISTING

DATE: 10/02/2002

PATENT APPLICATION: US/10/066,506A

TIME: 09:30:45

Input Set : N:\Cr4\09182002\J066506.raw

Output Set: N:\CRF4\10022002\J066506A.raw

```

1 <110> APPLICANT: Grogan, Case C.
2   Hevey, Michael C.
3   Schmaljohn, Alan, L.
4 <120> TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
5 <130> FILE REFERENCE: 003/243/SAP
6 <140> CURRENT APPLICATION NUMBER: US/10/066,506A
7 <141> CURRENT FILING DATE: 2002-01-31
8 <150> PRIOR APPLICATION NUMBER: 60/267,522
9 <151> PRIOR FILING DATE: 2001-01-31
10 <160> NUMBER OF SEQ ID NOS: 30
11 <170> SOFTWARE: Apple Macintosh Microsoft Word 6.0
13 <210> SEQ ID NO: 1
14 <211> LENGTH: 2252
15 <212> TYPE: DNA
16 <213> ORGANISM: Artificial Sequence
17 <220> FEATURE:
18 <223> OTHER INFORMATION: chimeric molecule between Ebola virus Zaire Mayinga strain
Glycoprotein 1
19   and Marburg virus strain Musoke Glycoprotein 2
20 <400> SEQUENCE: 1
21   atgggcgtta caggaatatt gcagttacct cgtgatcgat tcaagaggac   50
22   atcattcttt ctttgggtaa ttatcctttt ccaaagaaca ttttccatcc  100
23   cacttggagt catccacaat agcacattac aggttagtga tgtcgcacaaa  150
24   ctagtttgtc gtgacaaaact gtcatccaca aatcaattga gatcagttgg  200
25   actgaatctc gaagggaatg gagtggcaac tgacgtgcca tctgcaacta  250
26   aaagatgggg cttcaggtcc ggtgtccac caaaggtggt caattatgaa  300
27   gctggtgaat gggctgaaaa ctgctacaat cttgaaatca aaaaacctga  350
28   cgggagttag tgtctaccag cagcgccaga cgggattcgg ggcttcccc  400
29   ggtgccggtg tgtgcacaaa gtatcaggaa cgggaccgtg tgccgggagac  450
30   tttgcccttc ataaagaggg tgctttcttc ctgtatgata gacttgcttc  500
31   cacagttatc taccgaggaa cgactttcgc tgaaggtgtc gttgcatttc  550
32   tgatactgcc ccaagctaag aaggacttct tcagctcaca ccccttgaga  600
33   gagccggtca atgcaacgga ggaccctct agtggctact attctaccac  650
34   aattagatat caggctaccg gttttggaac caatgagaca gagtacttgt  700
35   tcgaggttga caatttgacc tacgtccaac ttgaatcaag attcacacca  750
36   cagtttctgc tccagctgaa tgagacaata tatacaagtg ggaaaaggag  800
37   caataccacg ggaaaactaa tttggaaggt caaccccgaa attgatacaa  850
38   caatcgggga gtgggccttc tgggaaacta aaaaaaacct cactagaaaa  900
39   attcgcagtg aagagttgtc tttcacagtt gtatcaaacg gagccaaaaa  950
40   catcagtggt cagagtccgg cgcgaacttc ttccgaccca gggaccaaca 1000
41   caacaactga agaccacaaa atcatggcct cagaaaattc ctctgcaatg 1050
42   gttcaagtgc acagtcaagg aagggaagct gcagtgtcgc atctaacaac 1100
43   ccttgccaca atctccacga gtccccaatc cctcacaacc aaaccaggtc 1150
44   cggacaacag cacccataat acaccctgtg ataaacttga catctctgag 1200

```

## RAW SEQUENCE LISTING

DATE: 10/02/2002

PATENT APPLICATION: US/10/066,506A

TIME: 09:30:45

Input Set : N:\Crf4\09182002\J066506.raw

Output Set: N:\CRF4\10022002\J066506A.raw

```

45 gcaactcaag ttgaacaaca tcaccgcaga acagacaacg acagcacagc 1250
46 ctccgacact ccctctgccg cgaccgcagc cggacccccca aaagcagaga 1300
47 acaccaacac gagcaagagc actgacttcc tggacccccgc caccacaaca 1350
48 agtccccaaa accacagcga gaccgctggc aacaacaaca ctcatcacca 1400
49 agataccgga gaagagagtg ccagcagcgg gaagctaggc ttaattacca 1450
50 atactattgc tggagtgcga ggactgatca caggcgggag aagaactcga 1500
51 cgatcgatcc tctggaggga aggcgacatg ttcccttttc tggatgggtt 1550
52 aataaatgct ccaattgatt ttgaccagcgt tccaaatata aaaacaatct 1600
53 ttgatgaatc ctctagtctt ggtgcctcgg ctgaggaaga tcaacatgcc 1650
54 tcccccaata ttagtttaac tttatcttat tttcctaata taaatgagaa 1700
55 cactgcctac tctggagaaa atgagaatga ttgtgatgca gagttaagaa 1750
56 tttggagcgt tcaggaggat gacctggccg cagggctcag ttggataccg 1800
57 ttttttggcc ctggaattga aggactttac actgctgttt taattaaaaa 1850
58 tcaaaacaat ttggtctgca ggttgaggcg tctagccaat caaactgccg 1900
59 aatccttgga actcttattg agagtcacaa ctgaggaaaag aacattctcc 1950
60 ttaatcaata gacatgctat tgactttcta ctacaagat ggggaggaac 2000
61 atgcaaagtg cttggacctg attgttgcgt cgggatagaa gacttgtcca 2050
62 aaaatatttc agagcaaatt gaccaaatta aaaaggacga acaaaaagag 2100
63 gggactgggt ggggtctggg tggtaaatgg tggacatccg actgggggtg 2150
64 tcttactaac ttgggcattt tgctactatt atccatagct gtcttgattg 2200
65 ctctatcctg tatttgctgt atctttacta aatatatcgg ataacggaat 2250
66 tc 2252

```

68 &lt;210&gt; SEQ ID NO: 2

69 &lt;211&gt; LENGTH: 747

70 &lt;212&gt; TYPE: PRT

71 &lt;213&gt; ORGANISM: Artificial Sequence

72 &lt;220&gt; FEATURE:

73 <223> OTHER INFORMATION: chimeric molecule between Ebola virus Zaire Mayinga strain  
Glycoprotein 1

74 and Marburg virus strain Musoke Glycoprotein 2

75 &lt;400&gt; SEQUENCE: 2

```

76 Met Gly Val Thr Gly Ile Leu Gln Leu Pro
77 1 5 10
78 Arg Asp Arg Phe Lys Arg Thr Ser Phe Phe
79 15 20
80 Leu Trp Val Ile Ile Leu Phe Gln Arg Thr
81 25 30
82 Phe Ser Ile Pro Leu Gly Val Ile His Asn
83 35 40
84 Ser Thr Leu Gln Val Ser Asp Val Asp Lys
85 45 50
86 Leu Val Cys Arg Asp Lys Leu Ser Ser Thr
87 55 60
88 Asn Gln Leu Arg Ser Val Gly Leu Asn Leu
89 65 70
90 Glu Gly Asn Gly Val Ala Thr Asp Val Pro
91 75 80
92 Ser Ala Thr Lys Arg Trp Gly Phe Arg Ser
93 85 90
94 Gly Val Pro Pro Lys Val Val Asn Tyr Glu

```

## RAW SEQUENCE LISTING

DATE: 10/02/2002

PATENT APPLICATION: US/10/066,506A

TIME: 09:30:45

Input Set : N:\Crf4\09182002\J066506.raw

Output Set: N:\CRF4\10022002\J066506A.raw

95		95		100
96	Ala Gly Glu Trp	Ala Glu Asn Cys Tyr Asn		
97		105		110
98	Leu Glu Ile Lys	Lys Pro Asp Gly Ser Glu		
99		115		120
100	Cys Leu Pro Ala	Ala Pro Asp Gly Ile Arg		
101		125		130
102	Gly Phe Pro Arg	Cys Arg Tyr Val His Lys		
103		135		140
104	Val Ser Gly Thr	Gly Pro Cys Ala Gly Asp		
105		145		150
106	Phe Ala Phe His	Lys Glu Gly Ala Phe Phe		
107		155		160
108	Leu Tyr Asp Arg	Leu Ala Ser Thr Val Ile		
109		165		170
110	Tyr Arg Gly Thr	Thr Phe Ala Glu Gly Val		
111		175		180
112	Val Ala Phe Leu	Ile Leu Pro Gln Ala Lys		
113		185		190
114	Lys Asp Phe Phe	Ser Ser His Pro Leu Arg		
115		195		200
116	Glu Pro Val Asn	Ala Thr Glu Asp Pro Ser		
117		205		210
118	Ser Gly Tyr Tyr	Ser Thr Thr Ile Arg Tyr		
119		215		220
120	Gln Ala Thr Gly	Phe Gly Thr Asn Glu Thr		
121		225		230
122	Glu Tyr Leu Phe	Glu Val Asp Asn Leu Thr		
123		235		240
124	Tyr Val Gln Leu	Glu Ser Arg Phe Thr Pro		
125		245		250
126	Gln Phe Leu Leu	Gln Leu Asn Glu Thr Ile		
127		255		260
128	Tyr Thr Ser Gly	Lys Arg Ser Asn Thr Thr		
129		265		270
130	Gly Lys Leu Ile	Trp Lys Val Asn Pro Glu		
131		275		280
132	Ile Asp Thr Thr	Ile Gly Glu Trp Ala Phe		
133		285		290
134	Trp Glu Thr Lys	Lys Asn Leu Thr Arg Lys		
135		295		300
136	Ile Arg Ser Glu	Glu Leu Ser Phe Thr Val		
137		305		310
138	Val Ser Asn Gly	Ala Lys Asn Ile Ser Gly		
139		315		320
140	Gln Ser Pro Ala	Arg Thr Ser Ser Asp Pro		
141		325		330
142	Gly Thr Asn Thr	Thr Thr Glu Asp His Lys		
143		335		340

## RAW SEQUENCE LISTING

DATE: 10/02/2002

PATENT APPLICATION: US/10/066,506A

TIME: 09:30:45

Input Set : N:\Crf4\09182002\J066506.raw

Output Set: N:\CRF4\10022002\J066506A.raw

144	Ile Met Ala Ser Glu Asn Ser Ser Ala Met	
145		345 350
146	Val Gln Val His Ser Gln Gly Arg Glu Ala	
147		355 360
148	Ala Val Ser His Leu Thr Thr Leu Ala Thr	
149		365 370
150	Ile Ser Thr Ser Pro Gln Ser Leu Thr Thr	
151		375 380
152	Lys Pro Gly Pro Asp Asn Ser Thr His Asn	
153		385 390
154	Thr Pro Val Tyr Lys Leu Asp Ile Ser Glu	
155		395 400
156	Ala Thr Gln Val Glu Gln His His Arg Arg	
157		405 410
158	Thr Asp Asn Asp Ser Thr Ala Ser Asp Thr	
159		415 420
160	Pro Ser Ala Thr Thr Ala Ala Gly Pro Pro	
161		425 430
162	Lys Ala Glu Asn Thr Asn Thr Ser Lys Ser	
163		435 440
164	Thr Asp Phe Leu Asp Pro Ala Thr Thr Thr	
165		445 450
166	Ser Pro Gln Asn His Ser Glu Thr Ala Gly	
167		455 460
168	Asn Asn Asn Thr His His Gln Asp Thr Gly	
169		465 470
170	Glu Glu Ser Ala Ser Ser Gly Lys Leu Gly	
171		475 480
172	Leu Ile Thr Asn Thr Ile Ala Gly Val Ala	
173		485 490
174	Gly Leu Ile Thr Gly Gly Arg Arg Thr Arg	
175		495 500
176	Arg Ser Ile Leu Trp Arg Glu Gly Asp Met	
177		505 510
178	Phe Pro Phe Leu Asp Gly Leu Ile Asn Ala	
179		515 520
180	Pro Ile Asp Phe Asp Pro Val Pro Asn Thr	
181		525 530
182	Lys Thr Ile Phe Asp Glu Ser Ser Ser Ser	
183		535 540
184	Gly Ala Ser Ala Glu Glu Asp Gln His Ala	
185		545 550
186	Ser Pro Asn Ile Ser Leu Thr Leu Ser Tyr	
187		555 560
188	Phe Pro Asn Ile Asn Glu Asn Thr Ala Tyr	
189		565 570
190	Ser Gly Glu Asn Glu Asn Asp Cys Asp Ala	
191		575 580
192	Glu Leu Arg Ile Trp Ser Val Gln Glu Asp	

## RAW SEQUENCE LISTING

DATE: 10/02/2002

PATENT APPLICATION: US/10/066,506A

TIME: 09:30:45

Input Set : N:\Crf4\09182002\J066506.raw

Output Set: N:\CRF4\10022002\J066506A.raw

```

193          585          590
194      Asp Leu Ala Ala Gly Leu Ser Trp Ile Pro
195          595          600
196      Phe Phe Gly Pro Gly Ile Glu Gly Leu Tyr
197          605          610
198      Thr Ala Val Leu Ile Lys Asn Gln Asn Asn
199          615          620
200      Leu Val Cys Arg Leu Arg Arg Leu Ala Asn
201          625          630
202      Gln Thr Ala Lys Ser Leu Glu Leu Leu Leu
203          635          640
204      Arg Val Thr Thr Glu Glu Arg Thr Phe Ser
205          645          650
206      Leu Ile Asn Arg His Ala Ile Asp Phe Leu
207          655          660
208      Leu Thr Arg Trp Gly Gly Thr Cys Lys Val
209          665          670
210      Leu Gly Pro Asp Cys Cys Ile Gly Ile Glu
211          675          680
212      Asp Leu Ser Lys Asn Ile Ser Glu Gln Ile
213          685          690
214      Asp Gln Ile Lys Lys Asp Glu Gln Lys Glu
215          695          700
216      Gly Thr Gly Trp Gly Leu Gly Gly Lys Trp
217          705          710
218      Trp Thr Ser Asp Trp Gly Val Leu Thr Asn
219          715          720
220      Leu Gly Ile Leu Leu Leu Leu Ser Ile Ala
221          725          730
222      Val Leu Ile Ala Leu Ser Cys Ile Cys Arg
223          735          740
224      Ile Phe Thr Lys Tyr Ile Gly
225          745

```

227 &lt;210&gt; SEQ ID NO: 3

228 &lt;211&gt; LENGTH: 1841

229 &lt;212&gt; TYPE: DNA

230 &lt;213&gt; ORGANISM: Artificial Sequence

231 &lt;220&gt; FEATURE:

232 <223> OTHER INFORMATION: chimeric molecule between Marburg virus strain Musoke  
Glycoprotein 1 and

233 Ebola virus Zaire Mayinga strain Glycoprotein 2

234 &lt;400&gt; SEQUENCE: 3

```

235      atgaagacca catgtttcct tatcagtctt atcttaattc aagggacaaa      50
236      aaatctcccc attttagaga tagctagtaa taatcaaccc caaaatgtgg      100
237      attcggtagt ctccggaact ctccagaaga cagaagacgt ccatctgatg      150
238      ggattcacac tgagtgggca aaaagttgct gattcccctt tggaggcatc      200
239      caagcgaatg gctttcagga caggtgtacc tccaagaat gttgagtaca      250
240      cagaggggga ggaagccaaa acatgctaca atataagtg aacggatccc      300
241      tctggaaaat ccttgctgtt agatcctcct accaacaatcc gtgactatcc      350
242      gaaatgcaaa actatccatc atattcaagg tcaaaaccct catgcacagg      400

```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 10/02/2002  
PATENT APPLICATION: US/10/066,506A      TIME: 09:30:46

Input Set : N:\Crf4\09182002\J066506.raw  
Output Set: N:\CRF4\10022002\J066506A.raw

**Invalid Line Length:**

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 18  
Seq#:2; Line(s) 73  
Seq#:3; Line(s) 232  
Seq#:4; Line(s) 278  
Seq#:5; Line(s) 409  
Seq#:6; Line(s) 459  
Seq#:7; Line(s) 605  
Seq#:8; Line(s) 655  
Seq#:9; Line(s) 801  
Seq#:10; Line(s) 852  
Seq#:11; Line(s) 998  
Seq#:12; Line(s) 1048  
Seq#:13; Line(s) 1193  
Seq#:14; Line(s) 1243

VERIFICATION SUMMARY

DATE: 10/02/2002

PATENT APPLICATION: US/10/066,506A

TIME: 09:30:46

Input Set : N:\Crf4\09182002\J066506.raw

Output Set: N:\CRF4\10022002\J066506A.raw